UNITED STATES UTILITY PATENT APPLICATION

FOR

A Method, System, and Apparatus

for a Credit based Flow control in a Computer System

Inventors:

Ling Cen

Prepared by:
Michael Nesheiwat
Patent Attorney

Intel Corporation
2111 N.E. 25th Avenue;
JF3-147
Hillsboro, OR 97124
Phone: (503) 712-8918
Facsimile: (503) 264-1729

Express Mail Number – EV325527958US

BACKGROUND

[0001] 1. Field

The present disclosure pertains to the field of protocols for computer systems. More particularly, the present disclosure pertains to a new method, system, and apparatus for an credit based flow control in a computer system..

[0002] Although the scope of the claimed subject matter is not limited in this respect, it is noted that some embodiments may include subject matter from the following co-pending applications: a first patent application with a serial number of , and with a Title of " **Simplified Two-Hop Protocol**", attorney docket P15925 and with the inventor Ling Cen.

[0003] 2. Description of Related Art

Computer systems may allow for connecting nodes with a Point to Point (PtP) network. Typically, a protocol controls the communication within a PtP network. For example, the protocol may be divided into a physical layer, link layer, and a protocol layer. The link layer may provide virtual channel (VC) services via a plurality of VC buffers.

[0004] Typically, the size of the VC buffers is designed such that there is no bubble or waste on a physical link when packets are sent between the nodes. Various factors affect the size of the link layer buffers, such as, round trip flight time, I/O delay, link buffer delay and multi-node arbitration delay. Thus, the link buffer's size is based on the average round trip delay from the transmission of a unit to the return of a credit. Therefore, each VC buffer for a physical link is sized according to the round trip delay (as exhibited in Figure 1). However, this is very inefficient. For example, a system would utilize 53,280 bits at each node for a typical configuration of five VCs per direction, 144

bit link unit, 60 entry receiver buffer and 10 entry transmit buffer. The number 53,280 was determined by adding the memory requirement for the receiver of [5 (# of VCs) x 64 x 144) to the transmitter of [5 (# of VCs) x 10 x 144.

Brief Description of the Figures

[0005] The present invention is illustrated by way of example and not limitation in the Figures of the accompanying drawings.

[0006] Figure 1 illustrates the prior art.

[0007] Figure 2 illustrates a schematic utilized in accordance with an embodiment.

[0008] Figure 3 illustrates a flowchart for a method utilized in accordance with an embodiment.

[0009] Figure 4 illustrates a chart utilized in accordance with an embodiment.

[0010] Figure 5 illustrates a system in accordance with one embodiment.

Detailed Description

[0011] The following description provides method and apparatus for improved multi-core processor performance despite power constraints. In the following description, numerous specific details are set forth in order to provide a more thorough understanding of the present invention. It will be appreciated, however, by one skilled in the art that the invention may be practiced without such specific details. Those of ordinary skill in the art, with the included descriptions, will be able to implement appropriate logic circuits without undue experimentation.

[0012] As previously described, a problem exists for protocols for a multi-node system. The prior art matches the number of VC's with the number of VC buffers and each of the VC buffer is sized according to the round trip delay. In contrast, the claimed subject matter efficiently utilizes VC buffers by incorporating at least one main buffer for each VC and sharing the remaining link buffers. The claimed subject matter may reduce memory size of the VC buffers and/or utilizes fewer bits in a packet header. The claimed subject matter will be discussed in further detail in Figures 2-5 of the specification.

[0013] Figure 2 illustrates a schematic 200 utilized in accordance with an embodiment. In one embodiment, schematic 200 illustrates a protocol for a link layer between two protocol layers for a multi-node computer system.

[0014] The schematic depicts an efficient link layer to facilitate transmitting and receiving data and information. In one embodiment, the link layer facilitates transmitting and receiving packets containing data, control information, and/or header information.

[0015] In one embodiment, the protocol and schematic allows the receiver to reserve one

link unit or one full packet for each VC, while the remaining link buffers are shared. For example, one VC allow for a Transmit (Tx) Link Buffer 102 and Receiver (Rx) Link 104. The remaining link buffers for the transmitter 108 and receiver 106 are shared according to a list buffer or FIFO (108 and 106) that is maintained for each VC. The list buffers are maintained for each VC. For example, the list buffer 108 is for the main transmit buffer and the list buffer 106 is for the main receiver buffer 104.

[0016] In one embodiment, the list buffer is to store the buffer indexes for all the link units for the respective VC. In the event that one list of link units for a particular VC is blocked, the protocol allows to switch to another VC's list of link units, which in essence is similar to switching from one VC's link buffer to another. Consequently, the VC is allowed to progress without an obstacle or blocking from another VC and a single main buffer is utilized.

[0017] In one embodiment, the memory size of main transmit or receiver buffer is based at least in part on: (the round trip delay) + (the number of VCs) – (1). In the same embodiment, the size of each main buffer applies to both the transmitter and receiver.

[0018] Therefore, the claimed subject matter facilitates a protocol that utilizes the following number of memory bits for each node:

[0019] Receiving : 5 (number of VCs) x (7)x(64 -4)x (size of list buffer)

[0020] Transmitting 5 (number of VCs)x(4)x(14) + 5(number of VCs)x(10+4)x(144),

[0021] Thus, the number of bits for the receiving and transmitting buffers results in 14,468 bits, a significant 4x drop in memory requirements as compared to the prior art.

[0022] However, the claimed subject matter is not limited to the previous example of memory requirements. One skilled in the art appreciates modifying the number of VCs or

number of shared link buffers or list buffer.

[0023] Figure 3 illustrates a flowchart for a method utilized in accordance with an embodiment. The method depicts one example of efficient memory usage for defining a link layer protocol. A block 302 depicts reserving one link unit or a packet for each virtual channel (VC). Subsequently, a block 304 depicts storing a plurality of buffer indexes of a plurality of link units in a link buffer or FIFO. Consequently, the remaining link buffers are shared based at least in part on the list buffer or the first in first out register (FIFO), as depicted by a block 306.

[0024] In one embodiment, an index from an entry of the list buffer or FIFO addresses a single main buffer for the single link unit. If one list is blocked for any reason, the link layer control switches to another virtual channel's list.

[0025] Figure 4 illustrates a chart utilized in accordance with an embodiment. In one embodiment, the chart illustrates a scheme for new credit passing encoding. For example, the sender utilizes a bit in the link unit for indicating whether this link unit is using the reserved credit or the shared VC buffer when a link unit of a particular VC is sent. In one embodiment, the sender initially utilizes the resources of the shared VC buffer and, when necessary, will utilize the reserved credit. Consequently, the reserved credit is not frequently used if the shared VC buffer is properly sized. Therefore, this result in saving bits that may be utilized for credit passing. For example, credits may be returned with a VC field and a credit field, as depicted in Figure 4. In one embodiment for a link buffer management, the shared VC credits are returned in a normal packet. In contrast, the reserved credits are returned in special link packets. Therefore, the bits in the header may be utilized for other purposes, such as, performance critical functions.

[0026] Figure 5 depicts a system in accordance with one embodiment. The system in one embodiment is a processor 502 that is coupled to a chipset 504 that is coupled to a memory 506. In one embodiment, the system is a point to point network of a multiple nodes for processors. For example, the chipset performs and facilitates various operations, such as, memory transactions between the processor and memory. In one embodiment, the system comprises one or all of the previous embodiments depicted in connection with Figures 2-4 of the specification to facilitate a credit based flow control.

[0027] While certain exemplary embodiments have been described and shown in the accompanying drawings, it is to be understood that such embodiments are merely illustrative of and not restrictive on the broad invention, and that this invention not be limited to the specific constructions and arrangements shown and described, since various other modifications may occur to those ordinarily skilled in the art upon studying this disclosure.